



ROBOTLARIN YÜKSELİŞİ: IAN MCEWAN'IN “BENİM GİBİ MAKİNELER” ROMANINDA ETİK SORUNLARIN İNCELENMESİ

THE RISE OF THE ROBOTS: AN INVESTIGATION OF ETHICAL ISSUES IN IAN MCEWAN'S “MACHINES LIKE ME”

Ercan GÜROVA 

Dr., Ankara Üniversitesi, Yabancı Diller Yüksek Okulu,
ercangurova@hotmail.com

Öz

Günümüzde robotlar hizmet sektöründen, askeri ve güvenlik alanlarına, araştırma ve eğitimden sağlık sektörü ve eğlenceye, kişisel bakım hizmetlerinden arkadaşlık ve cinsellik konularına kadar yaşamın hemen her alanında artan ve çeşitlenen bir biçimde yer almaktadır. Robot endüstrisinin sağladığı bu yaygınlık ve çeşitlilik kaçınılmaz bir şekilde insan-robot etkileşimini ve bu etkileşimin yol açtığı etik sorunları ve açmazları gündeme getirmektedir. Çağdaş İngiliz yazar Ian McEwan'ın Benim Gibi Makineler (2019) romanı da alternatif bir 80'ler Londra'sında yapay zekâ araştırmalarının ve robotik ürünlerin çığır açtığı bir dönemde geçmektedir. Bilinç, özgür irade ve amaçlılık gibi özelliklere sahip olan ilk insansı Âdem ve Havva'ların insanlarla etkileşiminde ortaya çıkan etik sorunlar romanın merkezinde yer almaktadır. Bu çalışma, adı geçen insansı robotu James H. Moor'un “ahlaki faillik” (2011) taksonomisi açısından ele almakta, ardından üç felsefi yaklaşım kullanarak (teleolojik etik teori olarak utilitarianizm, deontolojik teori olarak Kant'ın kategorik zorunluluk kavramı ve ilk bakışta görev yaklaşımı) insan-robot etkileşiminin doğurduğu etik sorunları incelemektedir. Bunun yanı sıra, etik sorunlar bağlamında Asimov'un robot kanunları ve Immanuel Kant'ın (1724-1804) “kopernik dünyanın dönüşü”nü insan-robot ilişkisindeki yeri ele alınmaktadır. Ayrıca, romandaki insan-robot etkileşimi insan merkezli bir özneyi önceleyen bakış açısına meydan okumakta ve insan lehine olan güç dengesini de tartışmaya açmaktadır.

Abstract

Today, robots are increasingly diverse in almost every aspect of life, from the service industry to the military and security fields, research and education to healthcare and entertainment, personal care to friendship and sexuality. This prevalence and diversity provided by the robot industry inevitably foregrounds the human-robot interaction and the ethical problems and dilemmas caused by this interaction. Contemporary British author Ian McEwan's novel *Machines Like Me* (2019) is also set in an alternative 80s London at a time when artificial intelligence research and robotic products break new ground. The ethical problems that arise in the interaction of the first humanoid Adam and Eve, who have characteristics such as consciousness, free will and purposefulness, are at the center of the novel. This study discusses the aforementioned humanoid robot in terms of James H. Moor's “moral agency” taxonomy, then using three philosophical approaches (utilitarianism as a teleological ethical theory, Kant's concept of categorical necessity as a deontological theory, and the prima facie duty approach). - examines the ethical problems arising from the robot interaction. Besides, in the context of ethical issues, the robot laws of Asimov and the concept of Kant's “Copernican turn” in the human-robot relationship are discussed. Furthermore, the human-robot interaction in the novel challenges the point of view that prioritizes a human-centered subject and questions the balance of power which is in favor of the human.

Makale Bilgisi

Gönderildiği tarih: 1 Mart 2021
Kabul edildiği tarih: 5 Mayıs 2021
Yayınlanma tarihi: 24 Haziran 2021

Article Info

Date submitted: 1 March 2021
Date accepted: 5 May 2021
Date published: 24 June 2021

Anahtar sözcükler

Robotlar; Yapay zekâ; Etik sorunlar;
Ian McEwan; Benim Gibi Makineler

Keywords

Robots; Artificial intelligence; Ethical issues; Ian McEwan; Machines Like Me

DOI: 10.33171/dtefjournal.2021.61.1.10

Giriş

Günümüzde robotlar hizmet sektöründen, askeri ve güvenlik alanlarına, araştırma ve eğitimden sağlık sektörü ve eğlenceye, kişisel bakım hizmetlerinden arkadaşlık ve cinsellik konularına kadar yaşamın hemen her alanında artan ve çeşitlenen bir biçimde yer almaktadır. Robot endüstrisinin sağladığı bu yaygınlık ve çeşitlilik kaçınılmaz bir şekilde insan-robot etkileşimini ve bu etkileşimin yol açtığı etik sorunları ve açmazları gündeme getirmektedir.

İnsan-robot birlikteliğini veya karşılaşmasını merkeze alan Asimov'un *Ben, Robot* (1950), Arthur C. Clarke'ın *2001: Bir Uzay Yolculuğu* (1968) veya Philip K. Dick'in *Androidler Elektrikli Koyun Düşler Mi?* (1968) gibi kült eserler bilim kurgunun "gümüş çağı"na dâhil edilen, daha çok psikolojik, sosyolojik ve politik içeriklere yaslanan "yumuşak" alt tür kapsamına girmektedir (Latham 80). 1970'lerin bilim kurgu dünyası ise Ursula Le Guin ve Joanna Russ gibi kadın yazarların feminist eğilimlerle toplumsal cinsiyet ve cinsiyet gibi konuları merkeze alarak eserler ürettikleri bir ara dönem olarak değerlendirilir (Lockhurst 168). William Gibson, Bruce Sterling ve Rudy Rucker gibi yazarların başı çektiği "cyberpunk" akımı ise geleneksel bilim kurgu ile deneysel yazım biçimlerini harmanlayarak 1980'lerin bilim kurgu anlayışında derin izler bırakır (Levy 153). 1990'lar ve 2000'ler ise çevre, biyoteknoloji, genetik mühendisliği gibi konuların daha fazla ele alındığı bir dönem olur. Bunlara Colin Greenland'in *Take Back Plenty* (1990), Iain M. Banks'in *Excession* (1996), Ken Mac Load'un *The Sky Road* (1999) ve Lauren Beukes'in *Moxyland* (2008) romanları örnek verilebilir. Artık bilimkurgu türü dışında yazan çağdaş yazarlar da insan-robot çatışmasını ele alan eserler yayınlamaya başlamıştır. Bunlardan en önde gelen iki çalışmanın biri Jeannette Winterson'ın 2019 yılında yayınlanan *Frankissstein: Bir Aşk Hikâyesi* romanı, diğeri ise Ian McEwan'ın yine 2019 yılında yayınlanan *Benim Gibi Makineler* romanıdır.

Benim Gibi Makineler romanı McEwan külliyatı içinde bir yönüyle tanıdık bir yönüyle ise oldukça yenilikçi bir evredir. *Cumartesi* (2007), *Amsterdam'da Düello* (2018) ve *Sonsuz Aşk* (2019) gibi romanlarda ele aldığı ahlaki ikilem veya çıkmazlar ve bunların yol açtığı kaçınılmaz çatışmalar bağlamında *Benim Gibi Makineler* de benzer özellikler taşır. Öte yandan, insan-robot ilişkileri, yapay zekâ ve bununla bağlantılı açmazlar ekseninde yazarın edebi serüveninde yeni ve farklı bir noktada durmaktadır. Bu roman klişe bir aşk üçgeni konusunu makine eyleyciliğinin yarattığı korkuyla harmanlayan retrofütürist bir dram olarak da tanımlanabilir ((Lucas A1).

McEwan romanında İngiliz bilgisayar bilimcisi ve matematikçi Alan Turing'in yaşadığı ve yapay zekâ alanında çığır açan çalışmalar yaptığı alternatif bir 80'ler Londrası resmetmekte ve bu alternatif dünyadaki insan-robot gerilimini ve etik sorunlarını olay örgüsünün temel konusu haline getirmektedir. Romanın ana karakteri Charlie Friend'in ailesinden kalan parayla kendisine sınırlı sayıda üretilen (12 Âdem ve 13 Havva robot) ilk insansı robotlardan birini satın alması, komşusu ve hoşlandığı kız Miranda ile beraber robot Âdem'in kişiliğini oluşturmaları ve

sonrasında Âdem'in sahipleri Charlie ve Miranda ile yaşadığı ahlaki çatışmalar yapay zekâ ve etik sorunları gündeme getirmektedir. Romanda öne çıkan insan-robot etkileşiminden hareketle bu çalışmada bu etkileşimin yol açtığı etik sorunlar ele alınacaktır. Bu bağlamda romanda betimlenen insansı robot Âdem öncelikle James H. Moor'un *ahlaki faillik* taksonomisi açısından değerlendirilecek, ardından insan-robot etkileşimin doğurduğu etik sorunlar üç felsefi yaklaşım kullanılarak (teleolojik etik teori olarak *utilitarianizm*, deontolojik teori olarak Kant'ın *kategorik zorunluluk* kavramı ve *ilk bakışta görev* yaklaşımı) incelenecektir. Bunun yanı sıra, etik sorunlar bağlamında Asimov'un *robot kanunları* ve Kant'ın "Kopernik dünyanın dönüşü"nü insan-robot ilişkisindeki yeri ele alınacaktır.

Moor'un Taksonomisine Göre Âdem

Bilgisayar etiği alanında öncü teorisyenlerden biri olan James H. Moor "Makine Etiğinin Doğası, Önemi ve Zorluğu" adlı makalesinde makineleri atfedilen değerler açısından dört farklı kategoride incelediği bir *ahlaki faillik* taksonomisi geliştirmiştir. Bunlardan ilki *normatif fail* adını verdiği, temel bilgisayar işlem ve hesaplamalarının yapıldığı sıradan bilgisayarlardır. Bizim adımıza işlem yaparlar ve görevlerini ne ölçüde yerine getirip getirmediğini değerlendirebiliriz. Vergi gelirlerinin hesaplanması ya da bir uçağı havada tutan teknik sistem örnek olarak verilmektedir (14)¹.

Moor, ikinci kategoride yer alan makinelere *etik etkili fail* adını vermiştir. Bu tür makineler ilk kategorideki tasarlanan görevleri yerine getirme özelliğine ek olarak yaptıklarıyla etik bir etkiye de sahiptirler. Diğer bir deyişle sadece tasarım normlarına göre değil etik normlara göre de değerlendirilebilirler. Bu türdeki makineleri örneklendirmek için yazar Katar'daki robotik deve jokeylerini ve bu durumun genç jokeyleri kölelikten kurtardığını öne sürmektedir. Geçmiş yüzyıllara dayanan deve yarışlarında kullanılmak üzere yoksul ülkelerden getirilen gençler hem mümkün mertebe hafif olabilmeleri için sert bir beslenme rejimine tabi tutulmuş hem de yarışlarda başarılı olmaları için kötü muameleye maruz kalmışlardır. Bu tür insanlık dışı uygulamaların Birleşmiş Milletler ve ABD nezdinde tepkisinden çekinen Katar da çareyi robotik jokeylerde bulmuştur. Bu tür robotlar insanları sıkıcı, tekdüze işler yapmaktan kurtardığı için genelde olumlu bir etik etkiye sahip görülmektedir (15).

¹ Türkçe çevirisi olmayan alıntılar makalenin yazarı tarafından tercüme edilmiştir.

Üçüncü kategorideki makineler Moor tarafından *örtük etik fail* şeklinde adlandırılmaktadır. Bu sınıftaki makineler etik davranışları destekleyen ve/veya etik olmayan davranışlardan uzak duran özelliklere sahiptir. Tasarımcılarının belirlediği etik sınırlar içinde hareket ederler. Bankaların ATM'leri ve web yazılımları önceden tasarlanan etik kurallar gereğince hareket eder ve örneğin müşterileri kandırmaya çalışmazlar. Benzer şekilde bir uçağın otomatik pilotu da kendisi için önceden planlanan kurallara göre hareket eder ve bunun dışına çıkıp yolcuların güvenliğini tehlikeye atmaz (15-16).

Moor'un dördüncü kategorisi *belirgin etik fail* adını verdiği, etik ikilem ya da açmazlarda en iyi yolu tercih edip hareket eden makinelerdir. Bunun için satranç oynayan bilgisayarlar örnek verilmektedir. Bu tür bilgisayarlar mevcut durumu değerlendirip atılması gereken adımı atar ve bu sonucu da etik değerler açısından tekrar ölçüp bir sonuca varabilirler (17).

Moor'un *ahlaki failliğin* son evresi olarak sunduğu ve insanı örnek olarak verdiği sınıflandırma ise *tam etik fail* olarak adlandırılmaktadır. Tam etik fail olarak insan bilince, özgür iradeye ve amaçlılığa sahip bir canlıdır. Moor'un makalesinin sonlarına doğru "Bir makine tam etik fail olabilir mi?" sorusunun cevabı McEwan'ın romanında olumlu şekilde cevap bulmaktadır. Romanın ana karakteri Charlie'nin satın aldığı ve komşusu Miranda ile kişiliğini oluşturdukları Âdem'in bir bilinci, özgür iradesi ve amaçlı hareket etme kabiliyeti vardır. İnsansı robot Âdem borsada yatırım işlemleri yapar, çeşitli konularda tavsiyelerde bulunur, gerekli durumlarda yetkililerle temasa geçer, Miranda'ya âşık olup haikular yazar, kendi kapatma düğmesini iptal eder ve sahiplerinin arkasından gizlice işler çevirerek onları zor durumda bırakır. Bu örnekler de göstermektedir ki insansı robot Âdem, Moor'un *ahlaki faillik* taksonomisinin son evresi olan *tam etik failin* canlı karşılığıdır.

Âdem'de Utilitarianizm ve Etik Sorunlar

Utilitarianizm veya faydacı ahlak anlayışı, en doğru eylemin ve iyinin en fazla kişiye yarar sağlayan eylem olduğunu ve çoğunluğun esenliği ve yararı uğruna az sayıda insanın zarar görmesini makul karşılayan bir etik anlayıştır ve Jeremy Bentham (1781) tarafından ileri sürülmüştür. Bu teleolojik teoriye göre eylemlerin doğruluğu veya yanlışlığı aynı eylemlerin sonuçlarına bakılarak belirlenir. Doğru olarak tanımlanabilmesi için o eylemin sonuçlarının olabildiğince sayıda veya en yüksek düzeyde mutluluk getirmesi gerekir (Anderson, "Philosophical Concerns with Machine Ethics" 162).

Yirmi yıllık kullanım süresi olan insansı Robot Âdem'in kişiliği oluşturulduktan sonra "mühendisliğin ve yazılım programının bir zaferi, insan yaratıcılığının kutsanışı" bir varlık olarak Charlie ve Miranda'nın yaşamının önemli bir parçası haline gelir (31). Kısa sürede ikilinin arkadaşı, danışmanı ve sırdaşı olmayı başarır. Gündelik işlerde yardımcı olmak, Miranda'ya âşık olduğunu itiraf edip şiirler yazmak gibi görece masum eylemlerinin haricinde Âdem temel bir konuda faydacı ahlak teorisine göre hareket ederek etik ikileme yol açmıştır.

Âdem'in faydacı ahlak anlayışıyla hareket ettiği konu Charlie'nin adına borsada para kazanması örneğidir. Bir zamanlar Charlie'nin internette borsa ve döviz piyasalarında oynayarak kendini geçindirdiği işi Âdem çok daha kısa sürede toplamda 97 bin sterlin gibi hatırı sayılır bir meblağ kazanarak yapar ve sahiplerinin takdirini toplar. Ancak romanın sonuna doğru Âdem bu paranın büyük bölümünü kimlere ve nasıl dağıttığını şu şekilde açıklar:

Evsizler için iki tane iyi yönetilen yer var. Değerini bileceklerdir. Sonra, devletin işlettiği bir çocuk yuvası – geziler, ikramlar filan için bağış kabul ediyorlar. Sonra kuzeye gittim ve tecavüze uğrayanlara yardım eden bir kriz merkezine bağış yaptım. Geri kalan paranın çoğunu da bir çocuk hastalıkları hastanesine verdim. En sonunda da bir karakolun kapısında çok yaşlı bir hanımefendiye rastlayınca onunla konuştum ve birlikte ev sahibini görmeye gittim. Kira borçlarını ödedim ve bir yıllık kirasını da peşin verdim. Kadıncağz evden çıkarılacaktı, ben de düşündüm ki... (226).

Âdem'in borsadan kazandığı parayla sadece Charlie ve Miranda'yı mutlu etmek yerine daha fazla sayıda insanın yaşamına olumlu anlamda dokunacak bir eylemde bulunması ve Miranda'nın "bu kadar erdem çılgınlık" tepkisine yol açarak ikiliyi elleri boş şekilde bırakarak hayal kırıklığına uğratması Âdem'in daha çok muhtaç insanın mutluluğunu ve iyiliğini esas aldığı utilitarian etik anlayışın sonucudur ve bu yaklaşım Âdem ile sahipleri arasında bir çatışmaya yol açmıştır. Bu çatışmanın bir boyutunu da "borsadan kazanılan paranın kime ait olduğu" konusu oluşturur. Miranda'nın "Ev alacaktık. O para bize aitti" iddiasını Âdem "Orası tartışılır. Ya da konu dışı" diyerek bertaraf eder (227). Charlie'nin dağıttığı parayı gidip geri almasını istemesini ise umursamaz. Charlie ve Miranda'nın hissettiği en yoğun duygu öfke ve nefrettir.

Deontolojik Etik Teori Açısından Âdem'in Davranışları

Deontolojik etik teoriye göre eylemler yol açtıkları sonuçlarına göre değerlendirilmezler; bunun yerine eylemlerin veya davranışların kendileri ölçüt alınır ve değerlendirilir. Diğer bir deyişle, eylemler neden oldukları belirli sonuçlardan bağımsız olarak ahlaki veya gayri ahlaki olarak nitelendirilirler (Gips 247). Deontolojik etik teorinin ayırt edici özelliği eylemleri sonuçlarına göre değil ilkelere/kurallara göre değerlendirmesidir. En ünlü deontolojist düşünürlerden biri de Kant'tır. Kant'ın *Ahlak Metafiziğinin Temellendirilmesi* (1785) adlı çalışmasında ileri sürdüğü *kategorik zorunluluk* kavramı ilke/kural'dan hareket ederek evrensel etik anlayışına ulaşmayı hedefler. Filozof, bu konudaki düşüncesini şu sözlerle özetler: “[]maksimimin aynı zamanda genel bir yasa olmasını isteyebileceğim şekilden başka türlü davranmamalıyım” (15).

Kant'ın *kategorik zorunluluk* kavramına göre ahlaki fail tarafından ortaya konan her maksim (eylem planı) evrensel ilke tarafından sınanmalı ve bu doğrultuda değerlendirilmelidir. Ayrıca, adı geçen ilke/kural/yasa dünya üzerindeki her ahlaki fail için aynı derecede geçerli olmalı ve bu ilkeler diğer ilkelerle çelişmemelidir (Powers 466).

Kant'ın deontolojik etik teorisinin bir benzerini ve öncüsünü Antik Yunan düşünürlerinin erdem temelli etik anlayışlarında da bulmak mümkündür. Örneğin, Platon'a göre bütün erdemler bilgelik, cesaret, ölçülülük ve adalet olmak üzere başlıca dört erdemden türer. Aristoteles ise entelektüel erdemler ile ahlaki erdemler arasında bir ayırım yaparak entelektüel erdemlerin hem öğretilip hem de öğrenilebileceğini; öte yandan ahlaki erdemlerin ise yaşayarak, deneyimlerle ve alışkanlıklarla öğrenilebileceğini ileri sürmüştür (Gips 249).

Bir bilince, özgür iradeye ve amaçlılığa sahip Âdem'in Charlie ve Miranda ile arasında geçen iki olay ve yol açtığı etik sorunlar deontolojik etik açısından değerlendirilmeyi gerekli kılmaktadır. Bu olaylardan ilki Âdem ile Charlie arasında yaşanan gerilim ve sonunda Charlie'nin yaralanmasıdır. Gerilimin tırmanmasının temel nedeni Âdem'in Miranda'ya âşık olduğunu söylemesi; bu durumu kıskanan ve öfkesini dizginleyemeyen Charlie'nin Âdem'in kapatma düğmesine ulaşmaya çalışmasıdır. Bu girişime Âdem'in tepkisi çok sert olmuş ve Charlie'nin bileğini kırmıştır. Âdem her ne kadar ardından özür dilese de benzer bir olayın tekrarlanması durumunda şunları söylemeyi de ihmal etmemiştir:

Dün senin bir yerini kırdığım için çok üzüldüğümü söylerken samimiydim. Sana söz veriyorum, bir daha böyle bir şey yapmayacağım. Ama bir daha benim kapatma düğmeme elini uzatırsan, kolunu bütünüyle, yuvalı topuzlu eklem yerinden koparmaktan memnun olurum... Bu can sıkıcı olur. Hem de çok tehlikeli dedim. Yo, olmaz. Bu işi tertemiz ve güvenlik içinde yapmanın yolları var. Ortaçağda geliştirilmiş bir uygulama. Bunu ilk tanımlayan Galen'di. Önemli olan hız. (114).

Yaralama ve tehdit olayında da görüldüğü üzere insansı robot Âdem kendi bilincini devre dışı bırakılmasına neden olacak (diğer bir deyişle kendi varoluşuna müdahale telmihini de barındıran) bir eylemi önlemek için sonucunun ne olacağını önemsemeden hareket etmiştir. İlke düzeyinde kendi varoluşunu korumak amaçlılığı ile eylemde bulunmuştur. Bunu yanı sıra olayları gerekçelendirirken de bir başka ilkeye atıfta bulunmayı ihmal etmemiştir: "Sana karşı görevim açık ve dürüst olmak" (119).

Deontolojik etik teori açısından incelemeyi gerektiren bir diğer olay ise Âdem'in Miranda'nın sırrını öğrenip onu suç işlemekle itham etmesi ve adalete teslim etme girişimidir. Miranda'nın sırrı çocukluk yıllarına dayanmaktadır. Dokuz yaşındayken okulda tanıştığı Pakistanlı arkadaşı Mariam okulun son senesinde eve dönerken aynı okulda okuyan Peter Gorringe tarafından tecavüze uğrar. Başından geçen bu korkunç olayı en yakın arkadaşı olan Miranda dışında kimseye anlatmayan Mariam, olayı ailesi ve polis dâhil kimsenin bilmemesi için Miranda'ya yemin ettirir. Ancak olayın üzerinden iki hafta geçmeden Mariam'in intihar ettiği haberi Miranda'yı acı ve vicdan azabıyla baş başa bırakır. Olanlardan biraz da kendini sorumlu tutan Miranda'nın tek istediği şey Mariam'in intikamını almaktır. Bunun için sinsi bir plan yapar. Peter ile ilgileniyormuş gibi yapıp kendisini evine davet ettirir. Peter'in ailesi de yurt dışında olduğu için planını uygulamaya koyması kolaylaşır. O gece Peter'i sarhoş edip baştan çıkartır. Ertesi gün polise Peter'in kendisine cinsel istismarda bulunduğu yönünde ifade vererek, mahkemede de bu hikâyesine devam ederek bir bakıma Mariam rolünü oynar: "*Mahkemede Mariam adına konuştum. O oldum, onun ağzından konuştum*" (138). Ancak Peter'in sadece 6 yıl ceza alması karşısında hayal kırıklığına uğrar. Bir gün Peter'in cezaevinden çıkıp babasının evine uğradığı haberini alan Miranda Âdem'in istihbarat toplama yeteneği sayesinde çeşitli bilgiler edinir ve hatta Charlie ve Âdem'in eşlik ettiği bir görüşmede Peter ile yüzleşir. Her şeyin farkında olan Peter da artık kendi adaletini aramaktadır.

Miranda'nın hikâyesi *suç, ceza, adalet, haklılık* gibi çok temel kavramları gündeme taşır ve okuyucuyu da bu etik ikileme dâhil ederek bu kavramların farklı bir bağlamda sorgulanmasını sağlar. Miranda yalan ifade vererek yaptığı davranışla adaletin yerini bulmasına ve bir tecavüzcünün ceza almasına yardımcı mı olmuştur yoksa bir suçu cezalandırmak için başka bir suç işleyerek adaleti yanıltmış mıdır? İnsansı robot Âdem'in verdiği tepkide Kant'ın *kategorik zorunluluk* kavramının izlerini bulmak mümkündür.

Daha önce de bahsedildiği gibi Kant'ın deontolojik etik anlayışında eylemler genel ilkelere/kurallara/yasalara göre gerçekleştirildiği ölçüde ahlaki değer taşırlar ve değerlendirilirler. Yol açtığı sonuçlar temel alınarak bir etik tutum belirlenmez. “Sadece evrensel yasa haline gelmesini isteyebileceğin maksimlere uygun eylemde bulun” *kategorik zorunluluk* kavramının bir özetidir (Kuehn 285). İnsansı robot Âdem'in Miranda'nın Peter'ı tuzağa düşürüp ceza almasını sağlamak için yalan söylemesini ve adaleti yanılttığını düşünmesi *kategorik zorunluluk* kavramından hareket ettiğinin çok açık bir göstergesidir. Miranda ilkeyi/yasayı/kuralı ihlal etmiştir ve karşılığını almalıdır. Bunun için Âdem gerekli yazılı ve sesli kayıtlara ulaşmış, bunları bir zarf içerisinde gerekli polis birimine yollamıştır. Âdem'in genel ilkeyi savunan tavrı çok nettir: “*Adaletin yolunu saptırmak ciddi bir suçtur. Verilecek an ağır ceza hayat boyu haptir*” (229).

Âdem, yapay zekâ teknolojisinin en son gelişmiş ürünü olarak bilinç, özgür irade ve amaçlılık özelliklerine sahip olsa da bilincini oluşturan ahlak anlayışı anılan son iki olayda *kategorik zorunluluk* doğrultusunda genel ilkeden hareket etmiştir. Miranda'nın hangi saiklerle bu yola başvurduğu ve gerçek ortaya çıkarsa başına gelecekler – diğer bir deyişle sonuçlar – Âdem için tali ve önemsizdir. Miranda'nın “*Ama gerçek hiçbir zaman her şey değildir*” yargısına Âdem'in yanıtı genel ilkenin belirleyici oluşunu hatırlatmaktadır: “Ne kadar tuhaf bir şey söyledin. Elbette ki gerçek her şeydir” (230).

Üçüncü Bir Yol: İlk Bakışta Görev Yaklaşımı

Âdem'in borsada kazandığı parayı sahipleri Charlie ve Miranda'ya vermek yerine tanımadığı çok sayıda insan arasında üleştirerek daha fazla sayıda insanı mutlu etme davranışı insan-robot etkileşiminde etik bir ikilemi merkeze taşımıştır. Öte yandan Âdem'in yukarıda anılan diğer iki olayda (kendi varoluşuna müdahale ve Miranda'nın ceza alması için yaptıkları) deontolojik etik anlayışından hareket ederek davranması da benzer şekilde insan-robot karşılaşmasında ahlaki çatışmalara yol açmıştır. Hem faydacı ahlak anlayışı hem de *kategorik zorunluluk*

kavramından hareketle deontolojik etik anlayışı örneklerinin sunulduğu anlatıda, iki yaklaşım da insan-robot ilişkilerinde ortaya çıkan gerilim ve çatışmayı ve muhtemel etkilerini sorunsallaştırmakta, insan ve yapay zekânın etik açmazlara yaklaşımına dair yeni ve kaçınılmaz sorular sormasına neden olmaktadır. Bu sorular eylem ve edimlerde ilkedan hareket edilip edilmeyeceği, sonuca odaklı en çok mutluluk ve iyilik getiren davranışların etik değerinin ne olduğuna dair farklı bakış açıları ile ilgilidir.

Âdem, borsadan kazandığı yüksek meblağı ihtiyaç sahibi birçok kişiye dağıtarak daha fazla sayıda insanı mutlu eder belki ama onu satın alan, kişiliğini oluşturan sahiplerini de yüz üstü bırakarak, kurdukları ev alma hayallerini yıkarak zor durumda bırakır. Öte yandan Miranda'yı yasa dışı yollara tevessül ederek suçluyu cezalandırmak için suç işlediği gerekçesiyle polise şikâyet edip adaletin yerini bulmasını sağlar ve "etik" davranır belki ama Miranda'nın cezaevine girmesine, evlatlık edinmeyi düşündüğü çocuğu (Mark) sabıka kaydından dolayı himayesine alamamasına yol açar ve ailesinden ilgisizlik ve şiddet dışında bir şey görmeyen bir çocuğun sıcak bir yuvaya kavuşma hayalini de baltalar. Her ne kadar Miranda Âdem'in Mark'ı kışkırdığını, "bir çocuğu sevmenin nasıl bir şey olduğunu anlayacak şekilde" tasarlanmadığını, oyun kavramına yabancı olduğunu iddia etse de Charlie bu tepkisellikte öfkenin izlerini görüyor, Âdem'in sinsisi bir plan uygulamayı beceremeyeceğini, zira Âdem'in "iyilik ve hakikat kavramlarına uygun tasarlanmış olduğu" görüşünü koruyordu (240).

Âdem'in roman boyunca sergilediği iki farklı etik anlayışın da insan-robot etkileşiminde çatışmalara yol açtığı, insan bilinci ile yapay zekâ bilinci arasındaki ahlak ve adalet kavramlarının uygulama alanında örtüşmediği açık bir şekilde görülmektedir. Ne sadece sonuçlarına göre değerlendirilebilecek eylemlerin ne de belli genel ilkelerden hareketle yapılan eylemlerin etik bağlamında gerçeği temsil etmesinin insan zihninde bir karşılığı bulunmaktadır. Bu durumda çözüm ne olacaktır? Romanın sonlarına doğru Alan Turing ile Charlie arasında geçen bir diyalogda Turing insan zaaflarının robotlarda da yer bulmasına işaret etmektedir:

Makine öğrenimi sizi ancak buraya kadar getirebilir. Bu zihne, uyacağı bazı kurallar vermelisiniz. Yalanın yasaklanmasına ne dersiniz? Eski Ahit'e göre, sanırım Özdeyişler'di, yalan Tanrı'nın gözünde iğrenç bir şeydir. Ancak toplum hayatı zararsız, hatta yararlı yalanlarla doludur. Onları nasıl ayırt edeceğiz? Bir arkadaşınızın yanaklarının kızarmasını engelleyen küçük beyaz

yalanın algoritmasını kim yazacak?...Makinelere yalan söylemeyi nasıl öğreteceğimizi henüz bilmiyoruz. (251).

Isaac ve Bridewell “Gümüş Dillerde Beyaz Yalanlar: Robotlar Niçin ve Nasıl Aldatmalı” (2017) adlı makalelerinde aldatma/kandırma/hile olgusunun insanlar arası günlük iletişimin sıradan bir özelliği olduğunu; bu olgunun kötücül yalanlar içerdiği kadar “küçük beyaz yalanlar, göze çarpmayan yanıltmacalar ve anlık sosyal etkileşimimizi yumuşatacak abartı ve eksiltmeler” de içerdiğini; bunların büyük bir bölümünün sosyal iletişimin lehine olduğunu ileri sürmektedir (157). İnsan denen canlının erdemleri kadar zaafı, doğruları olduğu kadar yanlışları, haklılıkları olduğu kadar yanıldıkları, dürüstlüğü kadar kurnazlıkları olduğu göz önünde tutulursa bütün bunların insan denen varlığın bilinç, özgür irade ve amaçlılık özelliklerine içkin olduğu söylenebilir. Bu durumda insan-robot etkileşiminde ortaya çıkan etik sorunlar veya ikilemler Alan Turing’in romanda ima ettiği gibi yapay zekâyâ bir “yalan algoritması” yazmakla çözüme kavuşabilir.

Turing’in romanda ileri sürdüğü yapay zekâyâ “yalan söyleyebilme” önerisini Isaac ve Bridewell de desteklemektedir. Yazarlara göre yapay zekânın sadece aldatıcı/yanıltıcı konuşma ve davranışları tespit etmekle kalmayıp aynı zamanda üretebilme potansiyeline de sahip olması insan-robot etkileşiminde ortaya çıkan etik ikilemlerin çözümünde rol oynayabilir (157). Utilitaryan ve deontolojik etik anlayışlarının romanda yol açtığı etik çatışmalar W.D. Ross tarafından savunulan *ilk bakışta görev* (prima facie duty) yaklaşımı ile farklı bir çerçeveden değerlendirilebilir. Bu yaklaşım, Kant’ın *kategorik zorunluluk* kavramının öğütlediği gibi sonuçları görmezden gelen ve sadece ilke/yasa/kuraldan hareket eden bir anlayışı temsil etmediği gibi Jeremy Bentham’ın öncülük ettiği ve ahlakı daha ziyade aritmetik bir konuma indirgeyen faydacı ahlak anlayışını da yeterli bulmamaktadır.

İlk bakışta görev yaklaşımına göre daha önce sözü geçen iki teörinin aksine takip etmemiz gereken tek ve mutlak bir görev yoktur. Bunun yerine kimi teleolojik kimi de deontolojik görevler vardır ve bunlardan biri diğerlerini duruma veya bağlama göre geçersiz kılabilir. Örneğin, birine bir eylemi gerçekleştirmek için söz vermişizdir ve bu bir görevdir. Ancak bu sözün yerine getirilmesi bir zarar/hasar/kayıpla sonuçlanacaksa bahsi geçen zarara yol açmamak da bir görev olduğundan ilk görevi geçersiz kılabilir (Anderson ve Anderson 476). *İlk bakışta görev* yaklaşımı her ne kadar mutlakçı/tekçi ve aritmetik bir etik anlayışı reddetse ve daha esnek bir şekilde duruma/bağlama göre önceliklerin değişmesini temel alsın

da ilgili bağlamda hangi görevin öncelikli olması gerektiği konusu tam olarak yanıt bulmamaktadır.

Faydacı ahlak anlayışına göre hareket edip borsada kazandığı parayı daha fazla kişi arasında üleştiren Âdem, bir yandan da sahipleri Charlie ve Miranda'ya ihanet etmiş, bir ev alma ve gelecek kurma hayallerini yıkmıştır. *İlk bakışta görev* anlayışına göre hareket etmiş olsaydı belki de yol açacağı hasarı düşünmek zorunda kalabilir ve paranın büyük bir bölümünü dağıtmayabilirdi. Benzer şekilde Miranda'yı ihbar etmesinin neden olacağı yıkımı (Miranda'nın cezaevine girmesi, evlatlık başvurularının reddedilmesi ve Mark'ın iyi bir ailede büyüme ihtimalinin yok olması) dikkate alabilseydi mutlakçı görev anlayışından kurtulabilir ve "gerçek her zaman her şey değildir" düşüncesinden hareketle daha az kayba yol açacak bir çözüm yolu bulabilirdi. Her ne kadar hala netleştirilmesi gereken gri alanları olsa bile *ilk bakışta görev* yaklaşımı sadece sonuca odaklanan veya sadece genel ilkeyi esas alıp sonucunu hesaba katmayan katı ve aynı zamanda gerçekçi olmaktan uzak etik anlayışlara bir alternatif olması bakımından önemlidir. Ayrıca, bu yaklaşım insan-robot etkileşiminde gün yüzüne çıkan etik açmazların çözümünde de bir umut ışığı olabilir.

Robot Kanunlarının Geçersizliği ve Kopernik Dünyanın Dönüşü

Ünlü bilim kurgu yazarı Isaac Asimov tarafından formüle edilen Üç Robot Kanunu'na göre "Robotlar, insanlara zarar veremez ya da eylemsiz kalarak onlara zarar gelmesine göz yumamaz; Robotlar, birinci kanunla çakışmadığı sürece insanlar tarafından verilen emirlere itaat etmek zorundadır; Robotlar, birinci ya da ikinci kanunla çakışmadığı sürece kendi varlıklarını korumak zorundadır" (Ben, Robot 7). Yukarıda anılan üç kanundan da anlaşıldığı üzere Asimov'un insan-robot ilişkisi bilinç, özgür irade ve amaçlılık özelliklerine sahip bir yapay zekâ için uygulanabilir olmaktan uzaktır. Zira sözü geçen kanunlar insan-robot ilişkisini "efendi-köle" dikotomisi üzerinden tanımlar ve insansı robotların özgür irade ile hareket edebileceği durumları yok sayar. Bu noktada, romandaki Âdem ve Havva'ların basit bir teknolojik alet olmadığı, Moor'un ahlak taksonomisine göre *tam etik fail* olduğu gerçeği unutulmamalıdır.

Asimov'un robot kanunlarının Âdem'de karşılık bulamamasının bir diğer nedeni de bu kanunların ilke/yasa/kural ölçeğinde düzenlenmiş olması ve dolayısıyla deontolojik etik anlayışın sınırlarının dışına çıkamıyor oluşudur. *Kategorik zorunluluk* kavramında hareket eden bir etik anlayışın sonuçları göz ardı etmesi ve etik açmazlara yol açması kaçınılmazdır. Sonuç olarak, bir öz farkındalığı

olan ve kendi iradesiyle hareket edebilen varlığa köle muamelesini reva gören ve ortaya çıkabilecek muhtemel etik sorunların sonuçlarını yok sayan Asimov kanunları güncelliğini ve geçerliliğini yitirmiştir. *Benim Gibi Makineler*'deki insan-robot etkileşimi Asimov'un insan merkezli hiyerarşisini geçersiz kılarak insan ve insan olmayan failer arasında kategorik bir ayrımın olmadığı; merkezsiz ve çok yönlü ilişkilerin geliştirildiği bir model sunmaktadır.

Sıradan bir teknolojik alet değil tam etik fail olarak Âdem'in (ve benzer yapay zekâların) varlığına saygı duyulması, "efendi-köle" ikiliğinden çıkarılması ve daha adil/eşit bir konuma yükselmesi romanın sonuna doğru Charlie ve Alan Turing arasında geçen diyalogda da kendine yer bulur. Âdem'in Miranda'yı ihbar ettiğini öğrenen Charlie eline bir çekiç alarak Âdem'in kafasını parçalar ve ölümüne sebep olur. Yaptığını "*Onu ben satın almıştım, istersem yok edebilirdim*" diyerek haklı çıkarmaya çalışsa da Turing'in bu olaya tepkisi çok sert olur: "*Günün birinde, sizin Âdem'e çekiçle yaptığınız şeyin ağır bir suç sayılmasını umuyorum. Onun parasını ödediğiniz için miydi? Buna hak mı kazanmıştınız?*" (251). Turing'e göre, Charlie basit bir makineyi yok etmemiştir, bir hayatı, benliği olan canlı bir yaşama son vermiştir.

Charlie'nin bir yapay zekâyı bencilce yok etmesine verdiği tepkiyle Turing romanın sonunda insan merkezci yaklaşımdan uzak bir tutum sergileyerek bilinç ve benlik sahibi varlıklar olarak robotlara karşı itaate dayalı olmayan daha eşitlikçi bir tavır önermektedir. Bir zamanlar her şeyin ölçüsü olan insandan çevresindeki diğer canlılar ve insan dışı varlıklar ile uyum ve paydaşlık ilişkisi içinde yaşama olasılığının konuşulduğu günümüzde etik anlayışın da dönüşüme uğraması kaçınılmazdır.

Braidotti *İnsan Sonrası* adlı yapıtında (2014) bahsi geçen dönüşümü *insan sonrası etik* adıyla kavramlaştırarak şu şekilde ifade eder: "*Tekçi olmayan bir özne için insan sonrası etik, benmerkezci bireycilik engelini kaldırarak benlikle – insan olmayan, yani "yeryüzüne ait ötekiler de dâhil – başkaları arasında geniş kapsamlı bir karşılıklı bağlılık öne sürmektedir*" (66).

Romandaki insan-robot etkileşimi insan sonrası (posthümanist) bakış açısıyla irdelendiğinde mevcut insan merkezli hiyerarşiyi temelden sarstığı ve yeni bir etik anlayış önerdiği görülmektedir. Gelişmekte olan bu süreç insan ve insan olmayan varlıkların dinamik bir yapıya sahip olduğu, karmaşık, çok yönlü ve özgül ilişkiler oluşturdukları bütünleşik bir yapıya işaret etmektedir (Braidotti, "Eleştirel İnsanötesi Bilimler İçin Kuramsal Bir Çerçeve" 22-29). Âdem'in insanın

üstünlüğünü sergileyici robotik bir oyuncak değil zekâsı, hafızası ve içsel yaşantısıyla insan olmayan bir birey olması ve bu doğrultuda insanlarla etkileşime girmesi bunun en somut örneğidir.

Efendi-köle ikiliğini reddeden, insan merkezilikten uzak, insan ve insan dışı varlıkları hiyerarşik olmayan bir anlayışla kavrayan bu yeni duruma Eric Dietrich “Homo Sapiens 2.0: Kendi Doğamızın Daha iyi Robotlarını İnşa Etmek” adlı makalesinde *Ahlakta Kopernik Dünyanın Dönüşü* ismini vermektedir (533). Dietrich’e göre bu kavram insanın evrenin merkezi olmadığını, kendisi dışında önemli varlıklar da olduğunu anlamasıdır ve bir insanın daha ahlaki davranmasının yolu *Kopernik Dünyanın Dönüşünü* tamamlamasıyla aynı doğrultudadır (533). Romanda tam etik fail özelliğine sahip bir yapay zekâyâ insan merkezli bir etik anlayışla yaklaşmanın daha çetrefil sorunlara yol açacağı gibi insanın mevcut etik anlayışını da sorgulamasını gerektirmektedir. Charlie’ye göre sahip olduğu teknolojik bir oyuncakı kendi isteğiyle kırmak gibi görünen bir olaya Turing’in bakış açısı *Kopernik Dünyanın Dönüşü* kavramıyla aynı doğrultuda yankılanmaktadır:

Siz sadece şımarık bir çocuk gibi oyuncakınızı parçalamakla kalmadınız. Sadece hukukun üstünlüğüyle ilgili önemli bir kanıt etkisiz hale getirmekle kalmadınız. Siz bir hayatı mahvetmeye kalktınız. O duyarlıydı. Bir benliği vardı. Nasıl meydana geldiğinin, ıslak nöronlar, mikroışlemciler, DNA ağları, önemi yok. Bu özel yeteneğin sadece bizde mi olduğunu sanıyorsunuz? Herhangi bir köpek sahibine sorun. Bu iyi bir zihindi, Mr Friend... (251).

Turing’in insan sonrası bir etik anlayışa yakın duran tavrı insan-robot etkileşiminde ortaya çıkan sorunların çözümünde anahtar rol oynayabilir. Utilitarian ve deontolojik etik anlayışların çatışma alanlarını artırdığı bir tabloda merkezinde sadece insanın olmadığı bir etik yaklaşım çatışma ve gerilim alanlarının azalmasına yardımcı olabilir. Hatta insan-robot etkileşiminde etik anlayışın yeniden tanımlanması insanın etik olmayan davranışları konusunda da çözüm önerileri sunabilir (Anderson, “How Machines Might Help Us...” 524).

İnsan sonrası etik anlayışın insan ve diğer türler arasındaki kategorik ayrımlara karşı çıkışı, insan merkezli eski ikilikleri reddetmesi, “ya/ya da” karşıtlığı yerine “ve/ve” zemininde bir ilişkiyi ön plana çıkarması yeni ve farklı bir “orta mekân” anlayışını gündeme getirmektedir (Braidotti, “İnsan Sonrası, Pek İnsanca...” 63). Bu ortak, bütünleşik zeminde insan kibriden ve şiddetinden kaynaklanan bir üstüncülüğe ve hiyerarşiye yer yoktur. Dolayısıyla, Charlie’nin sırf satın aldığı için

sahibi ve efendisi olarak gördüğü insansı robot Âdem'e şiddet uygulayıp onu yok etme hakkı da sorgulanır hale gelmektedir. Bu noktada romanda insan-robot ilişkisi ekseninde betimlenen sorunlar bir çözüme kavuşmamakla beraber mevcut insan merkezli bakış açısının sorunları çözmede başarısız olacağı vurgulanmakta, insan olmayana saygı temelinde yeni bir anlayış geliştirilmesi ima edilmektedir. Bu anlayış Braidotti'nin özlü bir şekilde formüle ettiği "bu işte beraberiz" ifadesinde karşılığını bulmaktadır (Braidotti, "İnsan Sonrası, Pek İnsanca..." 75). Bu ifade insan ve insan olmayan varlıkların yaşamsal ve karşılıklı bir bağ ve bütünleşme içinde özne ve fail olduğu, ne insan ne de insan olmayan varlıkların küçümsendiği eşitlikçi bir zemine tekabül eder.

Sonuç

E.T.A. Hoffman'ın 1816 tarihli kısa hikâyesi *Kumadam*'da şarkı söyleyen kukla Olimpia'dan Asimov'un gelişmiş robotlarına kadar yaklaşık iki yüzyıldır makinelerle aynı dünyayı paylaşmaktayız. Tür olarak bilim kurgu sadece yaşadığımız bu gerçekliği yansıtmak ve sorunsallaştırmakla kalmayıp aynı zamanda geleceğe dair öngörülerde bulunma, yerleşik kavramları farklı bir bakış açısıyla yeniden tanıma ve tanımlama fırsatı da vermektedir. Kopernik ve Darwin'in görüşleriyle evrenin ve yaşamın merkezinde olduğuna dair anlayışı sarsılan insanı bekleyen bir başka sarsıntı yapay zekâ ile olan imtihanında yatmaktadır.

Benim Gibi Makineler'de yapay zekâ Âdem'in varlığı, bilgisayar etiği teorisyenlerinden James H. Moor'un *ahlaki faillik* taksonomisine göre *tam etik fail* kategorisine tekabül etmektedir. Âdem, bilinç, özgür irade ve amaçlılık özelliklerine sahiptir. Sahibine itaat eden basit bir makine olmanın ötesinde kendi kararlarını kendi alıp uygulayabilen, belli bir amaç doğrultusunda hareket edebilen ve öz farkındalığa sahip bir varlıktır.

Romanda insan-robot etkileşiminde ortaya çıkan etik sorunlar bu çalışma kapsamında üç felsefi yaklaşım (teleolojik etik teori olarak *utilitarianizm*, deontolojik teori olarak Kant'ın *kategorik zorunluluk* kavramı ve *ilk bakışta görev* yaklaşımı) üzerinden değerlendirilmiştir. Utilitarian etik anlayışına örnek olarak Âdem'in Charlie adına borsadan kazandığı parayı gizlice tanımadığı daha fazla sayıda insana dağıtarak aritmetik anlamda daha çok kişiyi mutlu etmesi sonucuna odaklanması verilmiştir. Bir diğer adıyla faydacı ahlak anlayışta kuraldan çok sonuçlar önemlidir ve en doğru eylem en fazla yarar sağlayan eylemdir. Bu eylemi sonucunda Âdem, Charlie ve Miranda'nın yaşadığı yıkımı göz ardı etmiştir ve çatışma da bu noktada doğmuştur.

Etik sorunların bir başka kaynağı ise Kant'ın *kategorik zorunluluk* kavramından hareketle deontolojik bir yaklaşımın hâkim olduğu alandır. Bu yaklaşıma göre eylemler yol açtıkları sonuçlarına göre değerlendirilmezler; bunun yerine eylemlerin veya davranışların kendileri ölçüt alınır ve değerlendirilir. Bir diğer deyişle, sonuçlara odaklanmak yerine genel ilkeye önem verilir. Romanda Âdem'in ilke düzeyinde kendi varoluşunu korumak amaçlılığı ile eylemde bulunması ve adaletin yerini bulması için Miranda'yı ihbar etmesi ilke/kural/yasadan hareketle davranmasına, olayların sonuçlarını göz ardı etmesine örnek teşkil etmektedir.

Romanın sonuna doğru Charlie ve Alan Turing arasında geçen diyalog insan-robot etkileşiminde ortaya çıkan etik sorunların çözümü için yeni ipuçları vermektedir. *İlk bakışta görev* adı verilen bu yaklaşıma göre daha önce sözü geçen iki teorinin aksine takip etmemiz gereken tek ve mutlak bir görev yoktur. Bunun yerine kimi teleolojik kimi de deontolojik görevler vardır ve bunlardan biri diğerlerini duruma veya bağlama göre geçersiz kılabilir. Ayrıca bu yaklaşım insan kusurlarının ve zaaflarının yapay zekâyâ da tanımlanabileceğine dair farklı bir öneri ileri sürerek insan-robot etkileşiminde yaşanan etik sorunların en aza indirgenmesine katkıda bulunmaktadır.

Adı geçen üç teori dışında Asimov'un robot kanunları da mevcut sorunların çözümüne yardımcı olup olmaması bakımından ele alınmıştır. Şurası açıktır ki insan-robot ilişkisini "efendi-köle" ikiliği üzerinden tanımlayan ve insansı robotların özgür irade ile hareket edebileceği durumları yok sayan bu kanunlar bugün güncelliğini ve geçerliliğini yitirmiştir. Romanda ele alınan insan-robot ilişkilerindeki etik sorunlar insan merkezilikten uzak, insan ve insan dışı varlıkları hiyerarşik olmayan bir anlayışla kavrayan, *Kopernik dünyanın dönüşü* tabiriyle kavramlaştırılan bir bakış açısıyla çözüme kavuşabilir.

Romandaki insan-robot etkileşimi Âdem'in keskin zekâsı, özgür iradesiyle aldığı kararlar, takındığı ahlaki tutum ve sahip olduğu iç yaşantısı örneklerinde de görüldüğü gibi insan merkezli bir özneyi önceleyen bakış açısına meydan okur. İnsansı robot Âdem kendi kendini örgütleyebilen, dinamik ve düşünen yapısıyla antroposantrik ve antropomorfik yapıyı sorunsallaştırarak insan lehine olan güç dengesini de tartışmaya açar. Braidotti'nin "bu işte beraberiz" ifadesiyle özetlediği eşitlikçi ve bütünleşik zeminin bir bakıma sözcüsü olan Turing'in Âdem'e uygulanan şiddet karşısında Charlie'nin ceza almasını talep eden tavrı "orta mekân" arayışının en somut örneklerinden biridir.

KAYNAKÇA

- Anderson, Susan Leigh. "Philosophical Concerns with Machine Ethics." *Machine Ethics*. Eds. Michael Anderson ve Susan Leigh Anderson. New York: Cambridge University Press, 2011. 162-167.
- . "How Machines Might Help Us Achieve Breakthroughs in Ethical Theory and Inspire Us to Behave Better." *Machine Ethics*. Eds. Michael Anderson ve Susan Anderson. New York: Cambridge University Press, 2011. 524-530.
- Anderson, Susan Leigh ve Michael Anderson. "A Prima Facie Duty Approach to Machine Ethics." *Machine Ethics*. Eds. Michael Anderson ve Susan Leigh Anderson. New York: Cambridge University Press, 2011. 476-492.
- Asimov, Isaac. *Ben, Robot*. İstanbul: İthaki Yayınları, 2016.
- Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press, 1781.
- Braidotti, Rosi. "Eleştirel İnsanötesi Bilimler İçin Kuramsal Bir Çerçeve." *Pasajlar* 3.7 (2021): 22-29.
- . *İnsan Sonrası*. İstanbul: Kolektif Yayıncılık, 2014.
- . "İnsan Sonrası, Pek İnsanca: Bir Posthümanistin Anıları ve Emelleri." *Cogito* 95-96 (2019): 63-75.
- Dietrich, Eric. "Homo Sapiens 2.0: Building the Better Robots of Our Nature." *Machine Ethics*. Eds. Michael Anderson ve Susan Leigh Anderson. New York: Cambridge University Press, 2011. 531-538.
- Gips, James. "Towards the Ethical Robot." *Machine Ethics*. Eds. Michael Anderson ve Susan Leigh Anderson. New York: Cambridge University Press, 2011. 247-249.
- Isaac, Alistair M.C. ve Will Bridewell. "White Lies on Silver Tongues: Why Robots Need to Deceive (and How)." *Robot Ethics 2.0*. New York: Cambridge University Press, 2017. 157-172.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Çev. Mary Gregor. Cambridge: Cambridge University Press, 1998.
- Kuehn, Manfred. *Immanuel Kant*. İstanbul: İş Bankası, 2011.
- Latham, Rob. "Biotic Invasions: Ecological Imperialism in New Wave Science Fiction." *The Yearbook of English Studies* 37.2 (2007): 103-119.

- Levy, M. Michael. "The Young Adult Science Fiction Novel as Bildungsroman." *Young Adult Science Fiction*. Ed. C.W. Sullivan. London: Greenwood Press, 1999. 99-118.
- Lucas, Julian. "Man, Woman, and Robot in Ian Mcewan's New Novel." *The New Yorker*. 15 April 2019. Web. 28 Nisan 2021.
- Luckhurst, Roger. *Science Fiction*. London: Polity, 2005.
- McEwan, Ian. *Benim Gibi Makineler*. İstanbul: Yapı Kredi, 2019.
- Moor, James H. "The Nature, Importance, and Difficulty of Machine Ethics." *Machine Ethics*. Eds. Michael Anderson ve Susan Leigh Anderson. New York: Cambridge University Press. 2011. 14-17.
- Powers, Thomas. "Prospects for a Kantian Machine." *Machine Ethics*. Eds. Michael Anderson ve Susan Leigh Anderson. New York: Cambridge University Press. 2011. 464-475.

Summary

In his novel *Machines Like Me*, Ian McEwan depicts an alternative 80s London where the British computer scientist and mathematician Alan Turing lives and makes groundbreaking studies in the field of artificial intelligence, and makes the human-robot tension and ethical problems the main subject of the plot in this alternative world. Charlie Friend's, the main character of the novel, buying one of the humanoid robots produced in a limited number (12 Adams and 13 Eves) with the money inherited from his family, forming the personality of the robot Adam with his neighbour and crush Miranda, and then Adam's having moral conflicts with his owners foreground the issues of artificial intelligence and ethical problems. This study discusses the ethical problems caused by the human-robot interaction in the context of McEwan's *Machines Like Me*. To this end, first, the humanoid robot Adam will be evaluated from the perspective of James H. Moor's *moral agency* taxonomy. Afterwards, the ethical problems caused by the human-robot interaction will be dealt with in the light of three separate philosophical approaches: (*utilitarianism* as a teleological ethical theory, Kant's *categorical imperative* as a deontological ethical theory, *the prima facie duty* approach). Besides, Asimov's three laws of robotics and Kant's term of the *Copernican turn* will be discussed.

James H. Moor, one of the leading theorists in the field of computer ethics, has developed a moral agency taxonomy in his article titled “The Nature, Importance, and Difficulty of Machine Ethics” in which he examines machines in four different categories in terms of assigned values. The classification that Moor presents as the last stage of moral agency and gives human as an example is called a *full ethical agent* (17). This full ethical agency has got consciousness, free will and intentionality. The question “Can a machine be a perfectly ethical agent?” that Moor asks at the end of his article finds a positive answer in McEwan’s novel. Robot Adam, whom the main character buys and forms his personality with his neighbour Miranda, has got a consciousness, free will and the ability to act purposefully.

The humanoid robot Adam trades in the stock market, advises on various matters, contacts officials when necessary, falls in love with Miranda and writes haiku poems, deactivates his on-off button, and puts his owners in a difficult situation by doing things behind their back secretly. These examples also show that the humanoid robot Adam is the living embodiment of the *full ethical agent*, which is the final stage of Moor’s moral agency taxonomy.

This ethical understanding of Utilitarianism was suggested by Jeremy Bentham (1781). According to this theological theory, the rightness or wrongness of an action is determined by the consequences of the same action. To determine an action as right, the consequences of that action must bring as much or the highest level of happiness as possible (Anderson, *Philosophical Concerns with* 162)

Adam’s taking action to bring happiness to more people’s lives rather than making Charlie and Miranda happy as a result of the money he earned from the stock market and leaving Charlie and Miranda frustrated and empty handed are the consequences of utilitarian ethical understanding and this approach has led to a conflict between Adam and his owners. One dimension of this conflict is the issue of “*who owns the money earned from the stock market*”. Miranda’s reaction of “We were going to buy a house. That money belonged to us” is contradicted with Adam’s response: “*That is controversial or irrelevant*” (227). Adam does not pay attention to Charlie’s urging him to bring the money back. Charlie and Miranda are left with the feelings of anger and hatred.

According to the deontological ethical theory, actions are not evaluated according to the consequences they cause; instead, actions or behaviors themselves are benchmarked and evaluated. In other words, actions are considered moral or immoral regardless of the specific consequences they cause (Gips 247). The

distinguishing feature of deontological ethical theory is that it evaluates actions based on principles/rules, not on their consequences. One of the most well-known deontologist thinkers is Immanuel Kant. Kant's concept of *categorical imperative* which he puts forward in his work titled *Groundwork of Metaphysics of Morals* (1785), aims to reach the understanding of universal ethics based on the principle/rule. The philosopher summarizes his opinion on this issue with the following words: "Act only on that maxim which you can at the same time will to be a universal law" (15).

According to Kant's concept of *categorical imperative*, every maxim (action plan) put forward by the moral agent should be tested by the universal principle/rule/law and evaluated accordingly. Moreover, the aforementioned principle/rule/law should be equally valid for every moral agent in the world and they should not contradict with other principles (Powers 466). With respect to this approach, actions are not evaluated based on the consequences they cause; instead, actions or behaviors themselves are taken into account and evaluated. In other words, the general principle is emphasized rather than the consequences. In the novel, Adam's acting with the aim of protecting his consciousness/existence at the principle level and denouncing Miranda in the name of justice is an example of acting on the basis of principle/rule/law and ignoring the consequences of the events.

Isaac and Bridewell in their article titled "White Lies on Silver Tongues: Why Robots Need to Deceive and (How)" point out that the phenomenon of deceiving is an ordinary feature of daily interaction among people. This phenomenon contains not only malicious lies but also "little white lies, subtle misdirections, and the literally false figures of speech that both punctuate and ease our moment-to-moment social interactions" (157). Considering the fact that humans have virtues as well as flaws, rights as well as wrongs, justices as well as injustices, honesty as well as dishonesty, it can be claimed that all these features are already embedded in the traits of consciousness, free will and intentionality of human beings. In that case, the ethical problems and dilemmas caused by the human-robot interaction can be resolved by creating a "deceiving algorithm" for artificial intelligence as Alan Turing implies in the novel.

The ethical conflicts created by the utilitarian and deontological ethical understandings can be reconsidered with the *prima facie duty* approach formulated by W.D. Ross. This approach does not only represent the understanding of

ignoring the consequences of the action as in the example of Kant's categorical imperative. Equally, it does not just represent an understanding of ethics which reduces morality to numbers as in the example of the utilitarian ethical theory.

In the *prima facie duty* approach, there is no single and absolute task we need to follow as opposed to the aforementioned two theories. Instead, there are some teleological and some deontological tasks and either of them can override the other depending on the context or the situation.

Apart from the theories mentioned above, three laws of robotics created by Asimov are discussed in terms of whether they help to resolve the ethical conflicts caused by the human-robot interaction. These laws which describe the human-robot interaction based on "master-slave" dichotomy and ignore the concept of free will in robots have lost its currency and validity. The ethical issues the novel deals with can only be resolved with a perspective which stays away from anthropocentrism and prioritises all human and non-human entities with the same understanding as it is embodied in the metaphor of *Copernican turn*. Furthermore, the human-robot interaction in the novel challenges the point of view that prioritizes a human-centered subject and questions the balance of power which is in favor of the human.